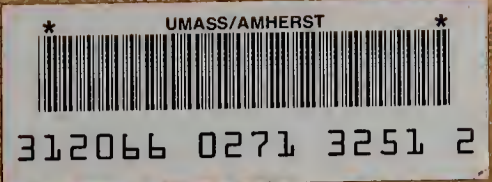


MASS. ED21.2: On 1/3/990/103



Grade Twelve

ON THEIR OWN

Students' Responses to Open-Ended Questions

In

Math, Reading, Science, Social Studies

EDUCATION DOCUMENT COLLECTION

1 1993

University of Massachusetts
Dipository Copy

Results of the 1990 Assessment

Massachusetts Department of Education

Massachusetts Educational Assessment Program

824/123

for individuals who can manage information, see patterns, identify needs, and solve problems. At the same time, those individuals who are most knowledgeable about the content itself have begun to re-examine what it means “to know” a discipline. In doing so, they are discovering common themes and concepts that underlie the various content areas and are suggesting that similar processes are involved in “coming to understand” any subject area.

A major consequence of these ideas is a shift in focus from learning as content knowledge *per se* to learning as the ability to use and interpret knowledge in critical and thoughtful ways. Subject matter has always held dominance in education. In elementary schools, the day is punctuated by shifts from reading to math to science to social studies, as students put away one set of books or papers and take out another. In middle schools and high schools, students move from class to class, subject to subject, without seeing the relevance of one subject to another. Even within subject areas, the layer-cake approach to curriculum obscures common ideas and themes, reinforcing the notion that subject area knowledge consists of a set of discrete facts and theories. However, the dominance of subject area knowledge is being challenged. As is discussed below, it is now argued that critical thinking is as relevant to literature as it is to science, social studies and mathematics; while problem solving is not the sole purview of mathematics, nor is hypothesis formulation limited to science.

This change in how we view the goals of education has implications for evaluation as well as instruction. If subject knowledge in itself is not a sufficient criterion for achievement, judgments of **correct** and **incorrect** in response to simple tests of skills and knowledge cannot be sufficient either. In order to measure how well a student performs, teachers have to be able to examine the process, not just the final product. Furthermore, individuals attempt to make sense of their perceptions and experiences, the associations that students make are often idiosyncratic and may be very different from the ones that were intended. These views of student learning demand a more open-ended form of testing, along with a more complex scheme of evaluation.

What Are Open-Ended Questions?

In defining open-ended questions, it may be easier to state what they are **not**. Open-ended questions are **not** multiple-choice questions without options. They are **not** questions that demand a single correct response. Nor are they questions in which

any response is acceptable. Rather, they are questions that address the essential concepts, processes and skills that go beyond the specifics of instruction to define the subject area. In general, they require complex thinking and yield multiple solutions. Unlike questions that can be judged right or wrong, they require interpretation and the use of multiple criteria on the part of the evaluator or teacher. Unlike questions that rely upon memorized facts, they demand thoughtfulness and a significant mental effort on the part of students.

Reading Background

During the past few years, the ideas that were expressed in *Reading and Thinking* have been refined and expanded by educators and philosophers interested in the reading process. Different themes have emerged from these activities.

One stresses the constantly shifting attitudes that the reader assumes in trying to understand any text. In contrast to the notion of reading as the accumulation of information, Judith Langer¹ proposes four different kinds of relationships that occur and reoccur during the process of reading.

Being out and stepping in. Readers use the information from the text and their background knowledge to get enough information to “step into” the author’s vision. In literature, readers try to make initial acquaintance with the character, plot and setting; while in exposition they try to figure out what the topic is about.

Being in and moving through. Readers immerse themselves in the author’s vision, trying to understand the meaning of the author. In exposition, readers take each new bit of information, try to understand it and to link it to what they already understand the text to have said about the topic. In fiction, readers use each new bit of information to go beyond what they already understand—asking questions about motivation, causality, and implications.

Being in and stepping out. Readers relate the text to their own knowledge and experiences. Readers of fiction use what they read in the text to reflect on their own lives, on the lives of others, or on the human condition in general. In non-fiction, readers use the text information to rethink information they already know.

Stepping out and going beyond. Readers distance themselves from the text and assume a critical stance, judging the text and relating it to other texts or experiences.

A second theme has suggested that literature is a powerful context for the teaching and learning of critical thought. As readers construct their understanding of the text, interpretations are often not possible. In fact, as Stephen Norris and Linda Phillips² suggest, the essence of critical reading is to raise alternative interpretations, weed out interpretations to the extent that available information will allow, and then remain with multiple possibilities. In their view, literary thinking is a complex reasoning process involving the analysis, synthesis, reformulation, linking, and generalization of ideas.

One important implication that can be drawn from this reappraisal of the reading process is that evaluation cannot be made purely on the basis of whether or not the reader's conclusions are similar to those of the teacher or test constructor. Instead, it is the quality of the reader's argument or justification that is of central importance.

Open-Ended Questions and the Results

(The actual questions, with scoring protocols and detailed results, are reported separately.)

Grade Four

Drawing inferences proved extremely difficult for fourth graders. Given a poem, students were asked to identify the underlying theme and to draw inferences about the characters. Although the context of the poem (children walking through the woods and finding objects) appeared to be appropriate for this age, the meaning of the poem depended upon an appreciation of different perceptions that people have. In this case, the poem contrasted the imaginative, fanciful vision of the poet to the more

prosaic and realistic vision of the other characters. Few of the children understood this. Among even the best readers (those at the 1550 and 1850 proficiency levels), only approximately 30 percent were able to state the underlying theme of the poem. When asked to use their interpretation to predict the action of the characters, the number dropped to 25 percent.

Students also experienced difficulty in interpreting the terms of an advertisement that was aimed at children of this age. Although 43 percent of fourth grade teachers believed that their students were well prepared to answer this type of question, this particular task was particularly challenging. Cluttered layout and deliberately ambiguous prose led the students astray. Despite the fact that the more proficient readers at this grade level showed a well-developed ability to handle details in multiple-choice questions, only approximately 30 percent were able to interpret the terms of this advertisement correctly. When asked to assume a critical stance by identifying specific examples of persuasive techniques, most students at all levels succeeded to some degree. However, only a few were able to characterize and describe the effects of such techniques on readers.

Grade Eight

The greatest discrepancy between teacher expectation and student performance occurred in response to a question which required students to identify and judge a set of assumption. Whereas less than 20 percent of teachers stated that their students were well prepared to answer this type of question, the overall success rate among students was more than twice that figure. On the other hand, performance did not follow the usual pattern of increased success rate with increased proficiency levels. Almost half of the most proficient students identified a valid assumption but, instead of judging the reasonableness of the assumption, they commented on the argument itself. Obviously, teachers did not view this question as one that fell within the area of reading instruction. They may have seen it as within the purview of social studies. Students also may have interpreted the question as one requiring argument rather than analysis. However, as discussed previously, the role of analytic

and evaluative skills are not confined to a specific subject area, but are part of any type of critical reading.

Grade Twelve

On the whole, the passages on which open-ended questions were based were longer and more demanding at grade 12 than at the preceding levels; however, similar abilities were tested.

Given a passage from the James Dickey's classic *Deliverance*, students were asked to infer the character of Lewis from description and dialogue. Sixteen percent effectively identified several traits or provided an in-depth discussion of one trait, while another 21 percent gave a less comprehensive response. There was a clear distinction in the quality of responses among the different proficiency levels. Forty-two percent of students who performed at the 1850 level in the multiple-choice section of the test gave an effective and convincing response, in contrast to 24 percent of students at the 1550 level, and many fewer at the lower levels.

The ability to infer meaning from text was tested in a practical context through the use of an article on teen-age traffic deaths. The article compared a number of disparate studies on drinking and driving and concluded that more stringent laws against drunk driving may be less effective than media coverage. Students were asked to identify the main point of the article and to justify their response. This proved to be challenging for even the most proficient readers. Only 21 percent of 1850 level and 12 percent at the 1550 level could give a clear and convincing argument.

The majority of students did not make the causal connection that was suggested by the article. Many stopped at the headline or made the general statements about drinking and driving without understanding the requirements for justifying their response. Few used the statistics in the article.

Implications for the Classroom

In large-scale testing programs the use of open-ended questions, although expensive and time-consuming to score, is justified by the amount of information they provide. Unlike short answer or multiple-choice questions, tasks that require

students to construct their own responses provide a window on students' thinking and understanding. Such tasks are vehicles for communicating actual achievement to parents, teachers, the public, and the students themselves. However, despite this obvious benefit, the most effective use of open-ended questions is in the classroom. Here, they provide models for students of the kinds of thinking that we wish to encourage. In addition, they give teachers the necessary information to improve their own effectiveness.

In developing their own open-ended questions, we offer teachers some general guidelines:

- ◆ Stress communication. Continually ask students to explain and to expand on their ideas, both in discussion and in written form. Let language become a vehicle for thought. Often, it is only through language that our thinking becomes clarified.
- ◆ Have students apply their skills in practical contexts. Set problems in the context of current affairs or the immediacy of everyday decisions. Not only will students become more motivated, you will help them realize the relevancy of their school learning and will encourage them to begin to transfer that knowledge to different contexts.
- ◆ Evaluate frequently. Testing encourages learning in at least two ways: it promotes review and consolidation; it highlights for the student what is of value to learn. Frequent testing also provides important information for the teacher. Not only does it help make instruction more focused, it provides evidence of students' understanding. In order to make valid and reliable judgments about levels of student attainment, we must use many different kinds of evidence in a range of contexts.

¹ Langer, Judith. (1989) *The Process of Understanding Literature*. Center for the Learning and Teaching of Literature, 1400 Washington Avenue, University at Albany, Albany, NY 1222.

² Norris, Stephen and Linda Phillips. (1987) "Explanations of reading comprehension: schema theory and critical thinking theory." *Teacher College Record*. Vol. 89, 2, pp. 282-306.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Science

Introduction

The Massachusetts Educational Assessment Program (MEAP) was established in 1985 in order to compare effectiveness among public schools and to give guidance for the improvement of curriculum and instruction. To accomplish this dual purpose, the Program has included in its assessment, not only what is currently taught in schools, but also the kinds of knowledge and thinking that educators believe should be taught if students are to be prepared to function well in a changing society. In addition to a broad range of multiple-choice items, the Program has included open-ended questions that require more than the identification of a correct option. Focusing on students' understanding, as well as their ability to reason and to apply knowledge in less traditional contexts, these open-ended questions communicate levels of student achievement more clearly than do multiple-choice items and give better guidance for instruction.

While in the 1990 assessment only 6 percent of all students were administered open-ended questions, beginning in the 1992 assessment, open-ended questions will be administered to all students and will contribute to school and district scores. Consequently, the Department is supplying the 1990 open-ended questions and results to schools in order to:

- ◆ Communicate levels of student achievement throughout the state.

- ◆ Familiarize teachers and administrators with the types of questions that will be included on the next assessment.
- ◆ Improve the assessment that takes place within the classroom by providing models which teachers can adapt in their own evaluation of students' knowledge, understanding, and abilities.

Rationale for Open-Ended Questions

During the past decade, we have witnessed significant shifts in the ways learning and instruction are viewed. Cognitive researchers have provided evidence about the complexity of learning. They have shown that conceptual understanding is more than the accumulation of knowledge, but depends on an active restructuring of old ideas to accommodate new experiences. These research findings, with their emphasis on personal accountability, have resonated in the practical world. As computers become repositories for information, both policy planners and businessmen have noted an increasing need

for individuals who can manage information, see patterns, identify needs, and solve problems. At the same time, those individuals who are most knowledgeable about the content itself have begun to re-examine what it means “to know” a discipline. In doing so, they are discovering common themes and concepts that underlie the various content areas and are suggesting that similar processes are involved in “coming to understand” any subject area.

A major consequence of these ideas is a shift in focus from learning as content knowledge *per se* to learning as the ability to use and interpret knowledge in critical and thoughtful ways. Subject matter has always held dominance in education. In elementary schools, the day is punctuated by shifts from reading to math to science to social studies, as students put away one set of books or papers and take out another. In middle schools and high schools, students move from class to class, subject to subject, without seeing the relevance of one subject to another. Even within subject areas, the layer-cake approach to curriculum obscures common ideas and themes, reinforcing the notion that subject area knowledge consists of a set of discrete facts and theories. However, the dominance of subject area knowledge is being challenged. As is discussed below, it is now argued that critical thinking is as relevant to literature as it is to science, social studies and mathematics; while problem solving is not the sole purview of mathematics, nor is hypothesis formulation limited to science.

This change in how we view the goals of education has implications for evaluation as well as instruction. If subject knowledge in itself is not a sufficient criterion for achievement, judgments of **correct** and **incorrect** in response to simple tests of skills and knowledge cannot be sufficient either. In order to measure how well a student performs, teachers have to be able to examine the process, not just the final product. Furthermore, individuals attempt to make sense of their perceptions and experiences, the associations that students make are often idiosyncratic and may be very different from the ones that were intended. These views of student learning demand a more open-ended form of testing, along with a more complex scheme of evaluation.

What Are Open-Ended Questions?

In defining open-ended questions, it may be easier to state what they are **not**. Open-ended questions are **not** multiple-choice questions without options. They are **not** questions that demand a single correct response. Nor are they questions in which

any response is acceptable. Rather, they are questions that address the essential concepts, processes and skills that go beyond the specifics of instruction to define the subject area. In general, they require complex thinking and yield multiple solutions. Unlike questions that can be judged right or wrong, they require interpretation and the use of multiple criteria on the part of the evaluator or teacher. Unlike questions that rely upon memorized facts, they demand thoughtfulness and a significant mental effort on the part of students.

Science Background

Currently, science curriculum is undergoing serious examination in light of recent findings indicating that United States students scored near the bottom in various international science exams. In response, studies (Science for All Americans, Research Within Reach: Science Education; 1994 NAEP Science Consensus) have generated curricular recommendations that are remarkably similar. Among those recommendations are:

Coverage should be in-depth.

By covering one topic thoroughly, students pick up skills that can be applied to other topics; they develop the habit of science. By stressing one concept and reinforcing that concept, students gain a depth of understanding than is far more substantial than a group of disparate facts. In other words, in the teaching of science, less is more.

Science should be taught as cross-discipline and multi-discipline.

Science is not a single area, but a series of related topics. Many of the concepts cross areas. The different subject areas do not exist in a vacuum and the inter-relationship of the various disciplines provide a totality that can be offered in no other manner. Particularly at the secondary level, science is taught as series of discrete disciplines. Nor is science isolated from other disciplines. The relationship between science and mathematics is obvious, but science and the

humanities, the fine arts, and the social sciences are also intrinsically tied to one another.

Science education should begin early.

Elementary teachers might be tempted to give short shrift to science because they may feel they don't have the training to teach science, yet early science education is vital to developing students who have a love and respect for science. We know that despite much instruction, students often retain misconceptions particularly about the natural world. The earlier teachers can begin to delve into these misconceptions, the earlier they can start correcting them.

Science should begin with questions, not with answers.

Instead of teaching facts to students, a more engaging way of teaching might be to ask "Why do you think...?" A major characteristic of scientists is curiosity. This approach nurtures a sense of curiosity and the resulting sense of accomplishment that accompanies it. Answers should not be definitive but rather lead into more and deeper questions.

Science should be done, not read or heard.

Science is a series of discoveries based on careful observation and calibration. How better to actively engage students in science than with a hands-on approach, particularly in the early years. Students who begin with hands-on science develop facilities for observation, handling equipment and developing hypotheses that can be transferred to other subject areas. Doing science is more than following a pre-selected experiment with orderly steps leading to a predetermined conclusion; it is observing phenomena, developing a hypothesis, testing that hypothesis, and replicating that test in order to confirm the results.

Science should be cooperative.

Since science in the real world is a cooperative activity, it makes sense that science in school should be taught in the same manner. A scientist does not work in isolation, but shares goals and data. The scientific community tends to share knowledge and resources for the greater good of the scientific world, hence the world at large.

Science should have a practical application.

Unless students can assign a personal value to science, it remains arcane and distant. Using science helps students to appreciate its relevance in their own lives. Knowing science can only help consumers and voters make educated decisions.

Open-Ended Questions and the Results

(The actual questions, with scoring protocols and detailed results, are reported separately.)

Grade Four

At grade four, students were asked to explain why or how something happened. In the area of earth science, they were asked to explain how water forms a canyon; in physical science, they were asked to explain condensation and evaporation. Among higher order thinking skills, students were asked to come up with a model of a natural phenomenon. They also had to complete the design of an experiment.

Teachers were given a selection of the open-ended questions prior to the test and asked to evaluate how well prepared their students were to answer the questions. At grade four, the relationship between actual performance on the questions and teachers' judgments of student preparedness for the questions varied. One-third of the teachers felt that their students were very well prepared to answer a question about evaporation and condensation, while eight percent of the students were able to discuss both evaporation and condensation successfully. On a question that asked students to consider characteristics of birds in relation to their habitats or food, 32% of the teachers felt that their students were very well prepared to answer. However, less than 15% of the students gave answers that were considered complete.

Grade Eight

The grade eight questions ranged from physics (laws of motion and friction and gravity) to physical science (the effects of temperature on a substance) to earth science (the causes of earthquakes) to scientific inquiry and process skills (steps in separating sand and sugar; reducing data in order to generalize;

generating an experiment). Other concepts included adaptation and the nature of sound waves.

As at grade four, teachers were given a small selection of questions and asked to evaluate their students' preparedness in responding to the task. Forty percent of the teachers felt that their students were very well prepared to describe an experiment, including the statement of hypothesis, design and necessary data. Approximately 15% of the students answered this question successfully. The largest discrepancy between teacher predictions and student performance was in response to the question "Explain why the moon does not fall to earth." Almost one-third of the teachers (31%) felt that their students were well prepared to answer this question, yet only 4% of the students could answer it successfully, with another 12% able to answer it with partial success.

Grade Twelve

At grade twelve, the questions focussed on interpreting data and making predictions including graphic representations. Other questions involved asking survey questions and generating a hypothesis. Students were also asked to explain why the moon doesn't fall to earth.

As in the case of the other grades, 12th grade science teachers were given the opportunity to consider the questions and respond to their students' preparedness. Given a scenario of fish dying unexpectedly, over one-fourth of the teachers predicted that their students were well prepared to generate two hypotheses and appropriate tests. About one-fourth of the students were able to generate good hypotheses and test them accordingly. Over one-third of teachers (35%) stated that their students were very well prepared to explain why the moon does not fall to earth, while 11% answered successfully. This question was problematic at both grades, with the teachers predicting much greater student preparedness than was the case. It may be that teachers felt they had covered the subject sufficiently.

At both grades eight and twelve, the majority of teachers (63% at grade 12 and 59% at grade 8) listed student indifference and lack of motivation as a very important reason why students do not perform well in science. And over 40% at both grades listed lack of preparation prior to entering the class as a cause of low achievement in science.

Implications for the Classroom

In large-scale testing programs the use of open-ended questions, although expensive and time-consuming to score, is justified by the amount of information they provide. Unlike short answer or multiple-choice questions, tasks that require students to construct their own responses provide a window on students' thinking and understanding. Such tasks are vehicles for communicating actual achievement to parents, teachers, the public, and the students themselves. However, despite this obvious benefit, the most effective use of open-ended questions is in the classroom. Here, they provide models for students of the kinds of thinking that we wish to encourage. In addition, they give teachers the necessary information to improve their own effectiveness.

In developing their own open-ended questions, we offer teachers some general guidelines:

- ◆ Stress communication. Continually ask students to explain and to expand on their ideas, both in discussion and in written form. Let language become a vehicle for thought. Often, it is only through language that our thinking becomes clarified.
- ◆ Have students apply their skills in practical contexts. Set problems in the context of current affairs or the immediacy of everyday decisions. Not only will students become more motivated, you will help them realize the relevancy of their school learning and will encourage them to begin to transfer that knowledge to different contexts.
- ◆ Evaluate frequently. Testing encourages learning in at least two ways: it promotes review and consolidation; it highlights for the student what is of value to learn. Frequent testing also provides important information for the teacher. Not only does it help make instruction more focused, it provides evidence of student's understanding. In order to make valid and reliable judgments about levels of student attainment, we must use many different kinds of evidence in a range of contexts.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Math

Introduction

The Massachusetts Educational Assessment Program (MEAP) was established in 1985 in order to compare effectiveness among public schools and to give guidance for the improvement of curriculum and instruction. To accomplish this dual purpose, the Program has included in its assessment, not only what is currently taught in schools, but also the kinds of knowledge and thinking that educators believe should be taught if students are to be prepared to function well in a changing society. In addition to a broad range of multiple-choice items, the Program has included open-ended questions that require more than the identification of a correct option. Focusing on students' understanding, as well as their ability to reason and to apply knowledge in less traditional contexts, these open-ended questions communicate levels of student achievement more clearly than do multiple-choice items and give better guidance for instruction.

While in the 1990 assessment only 6 percent of all students were administered open-ended questions, beginning in the 1992 assessment, open-ended questions will be administered to all students and will contribute to school and district scores. Consequently, the Department is supplying the 1990 open-ended questions and results to schools in order to:

- ◆ Communicate levels of student achievement throughout the state.

- ◆ Familiarize teachers and administrators with the types of questions that will be included on the next assessment.
- ◆ Improve the assessment that takes place within the classroom by providing models which teachers can adapt in their own evaluation of students' knowledge, understanding, and abilities.

Rationale for Open-Ended Questions

During the past decade, we have witnessed significant shifts in the ways learning and instruction are viewed. Cognitive researchers have provided evidence about the complexity of learning. They have shown that conceptual understanding is more than the accumulation of knowledge, but depends on an active restructuring of old ideas to accommodate new experiences. These research findings, with their emphasis on personal accountability, have resonated in the practical world. As computers become repositories for information, both policy planners and businessmen have noted an increasing need

for individuals who can manage information, see patterns, identify needs, and solve problems. At the same time, those individuals who are most knowledgeable about the content itself have begun to re-examine what it means “to know” a discipline. In doing so, they are discovering common themes and concepts that underlie the various content areas and are suggesting that similar processes are involved in “coming to understand” any subject area.

A major consequence of these ideas is a shift in focus from learning as content knowledge *per se* to learning as the ability to use and interpret knowledge in critical and thoughtful ways. Subject matter has always held dominance in education. In elementary schools, the day is punctuated by shifts from reading to math to science to social studies, as students put away one set of books or papers and take out another. In middle schools and high schools, students move from class to class, subject to subject, without seeing the relevance of one subject to another. Even within subject areas, the layer-cake approach to curriculum obscures common ideas and themes, reinforcing the notion that subject area knowledge consists of a set of discrete facts and theories. However, the dominance of subject area knowledge is being challenged. As is discussed in the Reading Overview, it is now argued that critical thinking is as relevant to literature as it is to science, social studies and mathematics; while problem solving is not the sole purview of mathematics, nor is hypothesis formulation limited to science.

This change in how we view the goals of education has implications for evaluation as well as instruction. If subject knowledge in itself is not a sufficient criterion for achievement, judgments of **correct** and **incorrect** in response to simple tests of skills and knowledge cannot be sufficient either. In order to measure how well a student performs, teachers have to be able to examine the process, not just the final product. Furthermore, as individuals attempt to make sense of their perceptions and experiences, the associations that students make are often idiosyncratic and may be very different from the ones that were intended. These views of student learning demand a more open-ended form of testing, along with a more complex scheme of evaluation.

What Are Open-Ended Questions?

In defining open-ended questions, it may be easier to state what they are **not**. Open-ended questions are **not** multiple-choice questions without options. They are **not** questions that demand a single correct response. Nor are they questions in which

any response is acceptable. Rather, they are questions that address the essential concepts, processes and skills that go beyond the specifics of instruction to define the subject area. In general, they require complex thinking and yield multiple solutions. Unlike questions that can be judged right or wrong, they require interpretation and the use of multiple criteria on the part of the evaluator or teacher. Unlike questions that rely upon memorized facts, they demand thoughtfulness and a significant mental effort on the part of students.

Math Background

In 1990 the National Council for Teachers of Mathematics published *The Curriculum and Evaluation Standards*, which challenged the traditional building block model of mathematics learning. According to this model, school mathematics consists of a set of discrete skills or building blocks—computational algorithms, measurement formulae, and algebraic rules. The elementary school teacher sets the first layer of blocks, the middle school teacher the second layer, and so on until, eventually the edifice is large and strong enough for students to be said to know mathematics. Although problem solving is a goal of learning, it is not actually connected to instruction. It comes later as a natural consequence of learning. The role of the teacher is to ensure that students solidify their learning at each level in order to progress to the next until, ultimately, they would be able to “do mathematics,” i.e., solve real problems.

Unfortunately, for most students, this model is not working. Students are not automatically solving problems. They are not making the connections. Furthermore, most are getting stuck at the lower levels, so they are not even allowed the chance to meet real problems.

In contrast, the Standards suggested a different model for mathematics learning—one of a network of skills and concepts that are related to one another. Students can go in and out of the network at any level because the ideas are connected. Furthermore, problems form

the context of this network, because solving problems is how we begin to understand what mathematics is about.

In their publication of the *Standards*, as well as the more recent *Professional Standards for Teaching Mathematics*, the National Council for Teachers of Mathematics stresses understanding over memory, relevance over training. It advocates an expansion of the curriculum to include new topics—such as probability and statistics, patterns and relationships, discrete mathematics—that reflect the mathematical needs of society. It gives greater emphasis to other topics—such as geometry and measurement—that have often been relegated to the back of the book. It delineates four basic abilities that are essential to all topics: reasoning, communication, problem solving, and the ability to see the connections between mathematical concepts. By including evaluation, it makes the point that assessment should be viewed as an integral part of instruction.

Open-Ended Questions and the Results

(The actual questions, with scoring protocols and detailed results, are reported separately.)

Grade Four

The open-ended questions were designed to cover different important concepts and abilities in the fourth grade mathematics curriculum. Questions concentrated on numerical operations and the ability to extract and use information presented in graphical form. There was little stress on calculation itself. Instead, students were asked to represent multiplication and division by writing stories and drawing pictures. They were required to interpret the data that were presented in a set of graphs and were asked to organize data. A final set of questions dealt with spatial visualization and geometric shapes.

When given a small selection of the actual questions, less than a quarter of the teachers replied that they believed their students to be well prepared to answer most; although another half believed that they were somewhat prepared. The exception was in response to a question asking students to write a story problem for a simple computation. In this case, 40 percent of the teachers replied that their students were well prepared.

In general, fourth graders were not hesitant to answer questions that may have been unfamiliar to them. In most cases, their answers showed a willingness to apply their reasoning and general knowledge to the task.

Grade Eight

The set of questions at this level contained both familiar word problems involving percentages and fractions and less common graphical problems that required students to extract and synthesize different types of information. As in the case of the fourth grade, teachers were presented with a selection of the problems and asked to respond in terms of their students preparation. It is not surprising that many teachers (between 45% and 71%) believed that their students were very well prepared to answer the word problems and those that probed students' understanding of numerical calculations. In contrast, only 20 percent responded that their students were prepared to use a map and a time/distance graph to describe a journey. Other questions—one concerning number patterns and one that asked students to synthesize and apply information from two graphs—were positively responded to by approximately a quarter of the teachers.

At this level, students' achievement reflected teaching more than in the earlier grade. Although few questions were answered well by more than 40 percent of the students, those tended to be the more familiar computational word problems. In response to these questions, high achievers on the multiple-choice section of the test were more successful than others. However, all students performed poorly on questions that asked them to observe, reason, and play with numbers. For example, there were few students who were able to draw inferences from graphical information or to identify more than one number pattern. In general, the eighth grade students were less prepared than their younger brothers and sisters to apply their reasoning in a mathematical context that was unfamiliar to them.

Grade Twelve

Again, at grade twelve, the preponderance of questions were in the area of numeration, with some questions focused on statistics and one which measured students' understanding of area. In most cases students were asked to explain or justify their response. Answers were judged on the quality of their response and credit was given for method in addition to the solution itself.

As was the case with the other grades, teachers were presented with a selection of questions and asked to comment in terms of their students' preparation. However, teachers who answered the twelfth grade questionnaire replied with reference to all grades at the high school level. With the sole exception of the value of 5!, no more than 20 percent felt that their students had been well prepared to answer any of this sample of questions. It seems that the written justification and reasoning required by these questions is not part of the usual high school curriculum.

Summary

From the teachers' point of view, students are not well prepared to answer questions that require them to synthesize information from various sources, to discover patterns, or to draw conclusions from data and argue their point of view. In general, the younger children—whether prepared or not—were more disposed to do so, and many performed well. As students progress through the grades, they appeared to be correspondingly less willing to attempt to answer questions that go beyond the specifics of the taught curriculum or required them to apply or discuss the concepts involved.

At grades 8 and 12 familiar questions that reflected the curriculum (e.g., the words problems at grade 8; questions that involved algebra at grade 12) were correlated with performance on the multiple-choice section of the test. The relationship between multiple-choice achievement and achievement with the less conventional questions was weaker, partially because few students at all proficiency levels performed well, partially because the explanations of the best students was not consistently better than those less capable.

Implications for the Classroom

In large-scale testing programs the use of open-ended questions, although expensive and time-consuming to score, is justified by the amount of information they provide. Unlike short answer or multiple-choice questions, tasks that require students to construct their own responses provide a window on students' thinking and understanding. Such tasks are vehicles for communicating actual achievement to parents, teachers, the public, and the students themselves. However, despite this obvious benefit, the most effective use of open-ended questions is in the classroom. Here, they provide models for students of the kinds of thinking that we wish to encourage. In addition, they give teachers the necessary information to improve their own effectiveness.

In developing their own open-ended questions, we offer teachers some general guidelines:

- ◆ Stress communication. Continually ask students to explain and to expand on their ideas, both in discussion and in written form. Let language become a vehicle for thought. Often, it is only through language that our thinking becomes clarified.
- ◆ Have students apply their skills in practical contexts. Set problems in the context of current affairs or the immediacy of everyday decisions. Not only will students become more motivated, you will help them realize the relevancy of their school learning and will encourage them to begin to transfer that knowledge to different contexts.
- ◆ Evaluate frequently. Testing encourages learning in at least two ways: it promotes review and consolidation; it highlights for the student what is of value to learn. Frequent testing also provides important information for the teacher. Not only does it help make instruction more focused, it provides evidence of student's understanding. In order to make valid and reliable judgments about levels of student attainment, we must use many different kinds of evidence in a range of contexts.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Social Studies

Introduction

The Massachusetts Educational Assessment Program (MEAP) was established in 1985 in order to compare effectiveness among public schools and to give guidance for the improvement of curriculum and instruction. To accomplish this dual purpose, the Program has included in its assessment, not only what is currently taught in schools, but also the kinds of knowledge and thinking that educators believe should be taught if students are to be prepared to function well in a changing society. In addition to a broad range of multiple-choice items, the Program has included open-ended questions that require more than the identification of a correct option. Focusing on students' understanding, as well as their ability to reason and to apply knowledge in less traditional contexts, these open-ended questions communicate levels of student achievement more clearly than do multiple-choice items and give better guidance for instruction.

While in the 1990 assessment only 6 percent of all students were administered open-ended questions, beginning in the 1992 assessment, open-ended questions will be administered to all students and will contribute to school and district scores. Consequently, the Department is supplying the 1990 open-ended questions and results to schools in order to:

- ◆ Communicate levels of student achievement throughout the state.

- ◆ Familiarize teachers and administrators with the types of questions that will be included on the next assessment.
- ◆ Improve the assessment that takes place within the classroom by providing models which teachers can adapt in their own evaluation of students' knowledge, understanding, and abilities.

Rationale for Open-Ended Questions

During the past decade, we have witnessed significant shifts in the ways learning and instruction are viewed. Cognitive researchers have provided evidence about the complexity of learning. They have shown that conceptual understanding is more than the accumulation of knowledge, but depends on an active restructuring of old ideas to accommodate new experiences. These research findings, with their emphasis on personal accountability, have resonated in the practical world. As computers become repositories for information, both policy planners and businessmen have noted an increasing need

for individuals who can manage information, see patterns, identify needs, and solve problems. At the same time, those individuals who are most knowledgeable about the content itself have begun to re-examine what it means “to know” a discipline. In doing so, they are discovering common themes and concepts that underlie the various content areas and are suggesting that similar processes are involved in “coming to understand” any subject area.

A major consequence of these ideas is a shift in focus from learning as content knowledge *per se* to learning as the ability to use and interpret knowledge in critical and thoughtful ways. Subject matter has always held dominance in education. In elementary schools, the day is punctuated by shifts from reading to math to science to social studies, as students put away one set of books or papers and take out another. In middle schools and high schools, students move from class to class, subject to subject, without seeing the relevance of one subject to another. Even within subject areas, the layer-cake approach to curriculum obscures common ideas and themes, reinforcing the notion that subject area knowledge consists of a set of discrete facts and theories. However, the dominance of subject area knowledge is being challenged. As is discussed in the Reading Overview, it is now argued that critical thinking is as relevant to literature as it is to science, social studies and mathematics; while problem solving is not the sole purview of mathematics, nor is hypothesis formulation limited to science.

This change in how we view the goals of education has implications for evaluation as well as instruction. If subject knowledge in itself is not a sufficient criterion for achievement, judgments of **correct** and **incorrect** in response to simple tests of skills and knowledge cannot be sufficient either. In order to measure how well a student performs, teachers have to be able to examine the process, not just the final product. Furthermore, as individuals attempt to make sense of their perceptions and experiences, the associations that students make are often idiosyncratic and may be very different from the ones that were intended. These views of student learning demand a more open-ended form of testing, along with a more complex scheme of evaluation.

What Are Open-Ended Questions?

In defining open-ended questions, it may be easier to state what they are **not**. Open-ended questions are **not** multiple-choice questions without options. They are **not** questions that demand a single correct response. Nor are they questions in which

any response is acceptable. Rather, they are questions that address the essential concepts, processes and skills that go beyond the specifics of instruction to define the subject area. In general, they require complex thinking and yield multiple solutions. Unlike questions that can be judged right or wrong, they require interpretation and the use of multiple criteria on the part of the evaluator or teacher. Unlike questions that rely upon memorized facts, they demand thoughtfulness and a significant mental effort on the part of students.

Social Studies Background

Currently, social studies is experiencing a “back to basics” movement stressing history, civics, and geography. This trend portends new curriculum examination and revitalization. Recent research has indicated that in elementary and middle grades, teachers make false assumptions about what their students know. Textbooks contribute to this problem by making the same assumption. In a recent article by Margaret McKeown and Isabel Beck,¹ researchers considered the background knowledge that middle school students bring to the study of history, the American Revolution in particular. They say that

“elementary social studies textbooks seem to assume knowledge of more specific concepts than students have under control...students do seem to have a general landscape involving the country being settled, the seeking of freedom and some notions of a process...Yet this seems exactly where the texts focus their material.”

This miscalculation of student knowledge calls for more complete and in-depth assessments of what significant knowledge is lacking. The longer that misconceptions continue uncorrected, the more difficult it becomes to correct them. Students will integrate knowledge to accommodate misinformation.

According to McKeown and Beck, one of the greatest challenges facing educators is finding and correcting

the confusions that young students bring to class. They suggest using a framework of the topic to organize previous knowledge and new information into a coherent whole. By presenting the relevant information in a framework, students would be better able to see the totality of the topic and evaluate whether their previous knowledge will fit in. They might be able to recognize irrelevant or inaccurate information using the framework as a sounding board.

Retention of knowledge is a problem at the secondary level. Social studies must build on the knowledge structures of the past. All this calls for a unified curriculum that considers how and when to teach materials. Walter Parker in *Renewing the Social Studies Curriculum*² suggests that themes in social studies can begin at the elementary grades and be reinforced across the years; that is, “content at any grade level should be presented in ways that provide, insofar as possible, a comprehensive view of a complex whole.” This suggests a thematic approach without sacrificing chronology or coverage, which would satisfy those who call for social studies to have breadth over depth.

Parker also discusses attributes of a successful social studies classroom.

- ◆ Lessons have substantive coherence and continuity. Disparate facts are pulled together.
- ◆ Students are given time to think—to prepare responses to questions.
- ◆ The teacher asks challenging questions and structures challenging tasks.
- ◆ Students are able to offer explanations and reasons for their conclusions.

Open-Ended Questions and the Results

(The actual questions, with scoring protocols and detailed results, are reported separately.)

Grade Four

At grade four, the open-ended questions covered process skills such as map reading and interpreting graphic representations. Another question focussed on a sense of history. Other questions involved conservation and economics.

Fourth grade teachers were given a small selection of these questions and asked to predict how prepared their students were to answer them successfully. Over half the teachers felt that their students were very well prepared to compare their own lives with those of colonial school children, while 33% of the students were able to do so. In response to how well their students were prepared to write a set of directions using a map, 31% of the teachers felt that their students were very well prepared. One-quarter of the students were able to do so.

Grade Eight

Eighth grade students were asked to interpret a political cartoon. A series of questions asked students to evaluate the quality of several arguments. Both specifics of history (students were asked to place six events on a time line) and a sense of history (students had to evaluate evidence to identify a speaker and when he lived) were covered. Students were asked to generate three courses of action and their negative impacts in response to an environmental crisis.

As in grade four, social studies teachers were given a small selection of open-ended questions and asked to evaluate how prepared their students were to answer those questions. About 15% of the teachers felt their students were very well prepared to generate three courses of action and their corresponding negative impacts; approximately 20% of the students were able to do so. When asked how well prepared their students were to explain why an argument was strong, about half the teachers felt that their students were very well prepared to do so. Eleven percent of the students were able to interpret and explain the strength of an argument.

Grade Twelve

Seniors were asked to interpret population graphs and make predictions based on those graphs. They were also asked to discuss how the policy of open immigration has affected the United States socially, politically, and economically. Another series of questions asked students to interpret two different graphs of immigration rates and discuss changes the graphs illustrated. As in grade eight, seniors were asked to interpret a political cartoon, this one about gun control.

Teachers who taught twelfth grade social studies were asked to comment on the preparedness of their students when given a sample of the questions. About one fourth (27%) of the teachers felt that their students were very well prepared to discuss the political, economic, and social effects of immigration while no more than 13% of the students were able to give complete answers in any of the categories. When given two graphs of immigration rates, one spanning 155 years, the other one year, 82% of the teachers felt their students would be able to answer why the numbers were larger in one graph than the other. Thirty-five percent of the seniors were able to answer that one graph spanned a much longer period of time than the other.

At grades eight and twelve, it appears that teachers felt their students were better prepared than they were to answer those questions that dealt with higher order thinking skills, such as evaluating arguments (grade eight) and cause and effect (effects of immigration at grade twelve).

In grades eight and twelve, the majority of teachers cited student indifference and lack of motivation as important factors in students' lack of success in social studies. Almost half the teachers at both grades cited poor preparation prior to entering the class as another important factor in students' lack of success.

Implications for the Classroom

In large-scale testing programs the use of open-ended questions, although expensive and time-consuming to score, is justified by the amount of information they provide. Unlike short answer or multiple-choice questions, tasks that require students to construct their own responses provide a window on students' thinking and understanding. Such tasks are vehicles for

communicating actual achievement to parents, teachers, the public, and the students themselves. However, despite this obvious benefit, the most effective use of open-ended questions is in the classroom. Here, they provide models for students of the kinds of thinking that we wish to encourage. In addition, they give teachers the necessary information to improve their own effectiveness.

In developing their own open-ended questions, we offer teachers some general guidelines:

- ◆ Stress communication. Continually ask students to explain and to expand on their ideas, both in discussion and in written form. Let language become a vehicle for thought. Often, it is only through language that our thinking becomes clarified.
- ◆ Have students apply their skills in practical contexts. Set problems in the context of current affairs or the immediacy of everyday decisions. Not only will students become more motivated, you will help them realize the relevancy of their school learning and will encourage them to begin to transfer that knowledge to different contexts.
- ◆ Evaluate frequently. Testing encourages learning in at least two ways: it promotes review and consolidation; it highlights for the student what is of value to learn. Frequent testing also provides important information for the teacher. Not only does it help make instruction more focused, it provides evidence of student's understanding. In order to make valid and reliable judgments about levels of student attainment, we must use many different kinds of evidence in a range of contexts.

¹ McKeown, Margaret G. And Beck, Isabel L. (Winter 1990). "The Assessment & Characterization of Young Learners' Knowledge of a Topic in History." *American Educational Research Journal*. Vol. 27, 4, pp. 688-726.

² Parker, Walter C. (1991). *Reviewing the Social Studies Curriculum*. Alexandria, Virginia: Association for Supervision and Curriculum Development.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Reading

Grade 12

Deliverance is a novel by James Dickey about four suburban men who go canoeing down a wild Georgia river in search of adventure and themselves. Read the following excerpt from *Deliverance* and answer the question that follows it.

Deliverance **September 14th**

The pale fire on the water was not subject to the current, and this seemed wonderful to me. It played and danced where it was, an invulnerable spirt that would die. We all sat without saying anything, and I was proud of us for that, and especially proud of Lewis, who I was afraid was going to expound. I stretched out on my back, paralleling the river.

There was a darkness on my inland side when I opened my eyes; I thought I had been lying there a long time. But then something filled the space again. It was Drew with his guitar. I sat up, and the water, though it still swarmed weightlessly with the cave-images of fire, now seemed on the point of swirling them down.

Drew tuned softly, then raked out a soft chord that flowed and floated away.

"I've always wanted to do this," he said. "Only I didn't know it."

He moved up the neck, drawing out chord after chord. These built and shimmered on each other in the darkness, in lonely harmony. Then he began to pick individual notes, and put the bass under them.

"It's woods music," he said. "Don't you think so?"

"Sure do."

I loved the powerful nasal country clang, the steely humming and the strings hit like hammers on rails. Drew played deep and clean, and neither of us could have been happier. He played "Expert Town" and "Lord Bateman"; he played "He Was a Friend of Mine" and "Shaggy Dad" and Leadbelly's "Easy, Mr. Tom."

"I really ought to have a twelve-string for this one," he explained, but it sounded good anyway.

Lewis brought over the cooked steaks while Drew played, and then we ate, two little steaks apiece and

big wedges of cake that Lewis' wife had made. The fire was leaving us; in the river it had already died.

"You know," Lewis said, "we don't have too many more years for this kind of thing."

"I guess not," I said. "But I can tell you, I'm glad we came. I'm glad to be here. I wouldn't be anywhere else, the way I feel."

"It's true, Lewis," Bobby said. "It's all true, what you said. It's great. And I think we did real good on the river. I mean, for amateurs."

"Yeah, good enough, I reckon," Lewis said. "But I'm sure glad you and I didn't get that sluggish wood canoe turned around backward just before we hit some white water. That might have been bad."

"We didn't though," Bobby said. "And I don't think it'll happen again, do you?"

"I hope not," Lewis said.

"Well, to the sleeping bags, men," I said stretching. . . .

I stood up, finally, and creaked and stooped into the tent. I was massively tired, and hated the laces of my tennis shoes which had hardened in the water until I couldn't untie them. I pulled the shoes off by main-strength, shucked off everything else as well and got into the bag and zipped it up. Drew was still playing, out on the bank; I could hear him trying out some high minor, far away. I lay back in the soft down, crinkling into the elastic resistance of the air mattress. I snapped out the flashlight and closed my eyes.

After reading the excerpt from *Deliverance*, what conclusions can you draw about the type of person Lewis is? Use specific details and evidence from the passage to support your response.

Major Concepts/Abilities Tested

Attention to specific details

Recognition of ambiguity

Ability to compare and contrast

Inference about character

Although Lewis plays a relatively small role in this excerpt from Dickey's novel, there are many clues to the type of person he is — his remarks, his actions, and the reactions that he provokes in others. Unlike the other characters, who seem to be relishing their day on the rapids, Lewis's remarks are begrudging, tinged with a note of fear. He admits that he is engaged on a rare adventure, but seems unwilling to savour it as he dwells on the disaster that might have befallen. There are other clues as well. The narrator is proud that Lewis refrained from "expounding," suggesting that there is a history of complaints. On the other hand, it is also Lewis who feeds his companions with steaks and a cake that his wife has made while they enjoy their rest. Students were graded on the extent to which they recognized these traits from the clues that were given.

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers

***** 16%** Thorough and complete responses that leave the reader with a distinct impression of the character of Lewis. These responses draw upon the entire excerpt to contrast Lewis's reactions to those of his companions. Their authors are able to cite accurate references to justify their conclusions.

"Lewis seems pessimistic and fatalistic, dwelling more on the negative than posi-

tive. He's the one that brings up the point that they are getting older and time and nature won't allow their bodies to be so adventurous anymore. He's also the person who brings up the fact that they almost capsized their canoe and when one of his friends tries to instill confidence in their future abilities and lift spirits by saying 'we didn't though and I don't think it will happen again.' Lewis refuses to give in to optimism by responding with an 'I hope not.' "

"Lewis seems to be pessimistic, almost fatalistic about life. His first line 'we don't have too many more years for this kind of stuff' shows how he thinks a lot about death. This idea is also suggested in his next line when he points out that they would have been in trouble if they had turned the canoe around. In contrast to the other men, who enjoy life, and are very carefree, Lewis worries about everything and doesn't enjoy the beauty and adventure of the outdoors."

"Lewis is the responsible and steadfast one in the group, preparing food for Drew and the speaker, but is also one for reflection and sentiment. 'We don't have too many years for this kind of thing' he says. Lewis is a realist and his sullen nature contrasts with the carefree, adventurous natures of the other two men. Lewis feels his mortality when challenged by the rapids in the river and is not ashamed to express his fear. While the other two are slightly more invincible and not as ready to express their fear. Lewis appreciates the company he is with and the atmosphere he is in, for he realizes that the time cannot last forever and should be treasured for as long as possible."

**** 11% Good inferences and accurate references to the text, but less developed than the previous responses. The reader gets a beginning sense of the character of Lewis but the response needs expansion or greater justification from the text.**

"Lewis seems to be nervous about the trip. He was afraid that the canoe was going to trip over when they hit the white water. He makes it seem like he's an old man because he says that they don't have many more years left to do that kind of thing anymore."

"Lewis seems like he worries a lot. He also seems like he thinks he is getting old. For example he says 'You know we don't have too many years for this kind of thing.' Lewis is kind of scared too when he mentions he is glad they didn't hit some white water because it could of been dangerous. Here shows examples of why Lewis worries a lot."

"Lewis is a person who is getting old because he says they don't have too many more years for this kind of thing. He is getting old and he isn't going to be able to camp anymore. He is also afraid of dying because he say he was glad that they didn't get the canoe turned around backward. If they did that could have resulted in something bad. I think that Lewis enjoys camping but he is afraid of some of the risks and hazards that it has."

*** 18% Minimal but correct interpretation, with little or no elaborations.**

"Lewis is a home type of person with a wife. He doesn't seem like he's been out in the wild before, but he seems to enjoy it now."

"He is realistic and takes things as they come but he is not an optimist. He has a negative outlook on things. You can tell this when they were talking about the boat ride about almost getting in some danger and that they don't have that many years of doing this kind of thing left."

***** good answer**

**** correct answer**

*** minimally correct answer**

Incorrect Answers

- 22% Inferences about Lewis's character that are based upon accurate but minor points in the text. By ignoring more pertinent facts to focus on less significant ones, the responses did not succeed in accurately characterizing Lewis.**

"Lewis is generous and adventuresome. He shows this by giving the men steaks from home and being an amateur white water rafter. Also Lewis is an older man. We can see this when he says 'You know, we don't have too many more years for this kind of thing.' "

"One can instantly observe after reading the passage that Lewis is an optimistic person. He gives forth a positive attitude. He is enthusiastic about the men and adventure. Lewis states 'You know we don't have too many years for this kind of thing.' The narrator displays his concern for Lewis in the first paragraph, he was 'especially proud of Lewis, who I was afraid was going to expound.' This shows that Lewis is a strong minded person, a determined person. Lewis also lets his friends know that he could go on with their adventure. Lewis offers support to Drew and the narrator."

"Lewis is a person that is sweet. He brought over steaks while Drew played. He's really nice. He has a wife."

- 25% Inferences based on incorrect reading of the text. Sometimes it is evident that the reader is confusing Lewis with the narrator.**

"Lewis seems like an adventurous person. He likes canoeing, and being out in nature. Camping, hiking, backpacking, things like that. He seems to worry about what's going to happen in the future. He seems grateful to be where he is, when he's there. He doesn't dwell on the past."

"Lewis is a calm, quiet down to earth person. He appreciates nature and being alive. He has a very positive look on life. The whole passage is full of descriptive details. I know that Lewis is an optimistic kind of person because of things like 'But I can tell you, I'm glad we came. I'm glad to be here. I wouldn't be any where else, the way I feel.' "

"Lewis seems to be a naturalist to me. He loves living in nature and living on the river. He said, 'He loved the powerful nasal country clang, the steely humming and the strings hit like hammers on nails.' 'We don't have too many more years for this kind of thing.' there it shows that he enjoys it and really doesn't want it to end. They had a fire set up with the tent and sleeping bags right on the side of the white river. His closing sentence, 'I lay back in the soft down, crinkling into the elastic resistance of the air mattress. I snapped out the flashlight and closed my eyes.' He seems to really enjoy living in the great outdoors."

- 5% Blank**

Comments

One of the great benefits of fiction is its ability to enrich our understanding of others. Through the careful selection of dialogue, tone, and description, the novelist provides the reader with opportunities to examine the attitudes and perceptions of other human beings in the kind of detail that is impossible in everyday life. In keeping with the notion of reading as a way of making sense of the world, it is important for students to recognize the clues that authors provide in establishing the personality and motivation of their characters.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Reading

Grade 12

*Read the following preface to the book **Vietnam: There and Here**; then answer question 3.*

Vietnam: There and Here — Preface —

Twenty-five years ago, many Americans had difficulty locating the country on a map. Today, many Americans have difficulty locating the war in their minds.

The Vietnam War was one of the most painful events in all of American history. In 1965, when the first American combat soldiers went to war in Vietnam, Americans went to war at home. Proponents of the war — called hawks — believed that the United States, born of the desire for liberty, cannot sit by when communism denies people their liberty. Opponents of the war — called doves — protested against the war for a variety of reasons. Some believed that Vietnam was torn by a revolution or a civil war and that it is not proper for the United States to meddle in the internal affairs of any country. Others believed the war was simply unwinnable.

When American troops finally came home, in 1973, they sometimes found themselves still embattled. Veterans were

often vilified by both proponents and opponents of the war, who saw them either as agents of American defeat or as agents of American immorality. In many cases the veterans were simply ignored. There were no heroes. Now, years after the war in Vietnam was ended, the war at home continues.

Many Americans have never come to a conclusion about the U.S. involvement in Vietnam. This seems strange, for never before had Americans at home been offered so much information about a war. Day after day, on television they saw footage of bombing runs, platoons moving slowly through jungles, body bags. In newspapers and magazines, they read heavily illustrated articles about corruption in the South Vietnamese government, the Tet Offensive, Vietcong assassinations. At the same time, the media were covering the war at home. Day after day, Americans saw and read about peace marches, draft dodgers, draft-card burnings. It was not long before many Americans came to regard protesters as the real enemy.

After the war, Americans still disagreed, but they did so in silence; Vietnam seemed almost a taboo subject. The renewal of discussion about the war is a sign of our health and our hope. It is clear that the Vietnam War will haunt us until we understand it. It is also clear

that the United States will be a weakened giant until we develop a new sense of when and how we will wage war. Too young to have heard the debate while the war was raging in Vietnam, you now hear the debate that has revived at home. This book chronicles the war in Vietnam and the turmoil it caused at home, and outlines the larger issues involved in both. I leave to you, to reader, to decide if — and why — the war was right or wrong. You were not a part of the war in Vietnam, but you will be a part of the new debate at home.

If you were to write the five chapters of a book with the title and preface above, what would five good chapter titles be?

Major Concepts/Abilities Tested

Understanding of text

Understanding of formal features such as format

Ability to recognize logical sequence

Responses

Although calling for many of the same logical skills as outlining, this task is more difficult. The purpose of a preface is to give an impression of what will lie ahead for the reader, not to lay out the content in a logical order. As a result, students had to use the author's description to speculate on the range of topics that would probably be covered and the ways in which they could best be described. They were also required to suggest the essence of each chapter in a succinct and, possibly, intriguing manner. Finally, they had to extract the logical structure of the book that the author was describing. Each of these three dimensions—range of topics, appropriateness of language, and logical sequence—was rated separately. Although success on the dimensions were strongly related to each other, students generally found it easier to express chapter titles in a succinct and intriguing manner than to structure them in a logical sequence.

Results

	range of topics	appropriate language for title	logical sequence
A best	21%	27%	29%
B average	44%	50%	35%
C poor	31%	18%	32%

no response: 4%

Below are typical student responses that reflect different ratings on the three dimensions.

Chapter 1: Vietnam: a General Overview

Chapter 2: Controversy Over the Vietnam War

Chapter 3: Heros or Not? Veterans of Vietnam

Chapter 4: Image of Vietnam at Home

Chapter 5: A Time For Reflection: The Impact of the Vietnam War

◇ ◇ ◇

Chapter 1: The Drafting

Chapter 2: Birds at War

Chapter 3: Assassinations and Snapshots

Chapter 4: When They Came Home, (No One Wa There)

Chapter 5: The War is Over (Or Is It?)

Both sets of responses capture the main points of the preface in a succinct and engaging way although the tone differs. The language in the first set suggests a more issues-oriented approach, while the second set is more colorful. The student uses word play and provocative phrases to capture the reader's attention. Both sets cover the range of possible topics that is suggested by the author of the preface and both are structured in a logical sequence.

Both sets of responses were rated as A on all three dimensions.

Chapter 1: The Reasons Why the Americans Became the Real Enemy

Chapter 2: The Hawks vs. the Doves

Chapter 3: The Answer to Why the U.S. Meddles in Other Countries Affairs

Chapter 4: Veterans and How They Were Treated After the War

Chapter 5: The Coverage of the Vietnam War

Although this response covers some of the main points that are mentioned in the preface, it neglects others such as the aftermath of the war. Language is also somewhat weak. Not only do some of the chapter titles lack the brevity and focus that is their characteristic format, they are also ambiguous (e.g. Chapters 1 and 3). Finally, possibly because of the ambiguity, the logic of the chapter sequence appears to be weak.

This response was rated as B on all three dimensions.

Chapter 1: Vietnam Will Never End

Chapter 2: The Opinionated Conflict

Chapter 3: The Rage That Lives Within

Chapter 4: Period of Realization

Chapter 5: Understanding Their Reasons

Here, the student appears to have ignored a great deal of the content that was referred to in the preface, namely, the cause of the war, the role of the media, and the treatment of the veterans. Although the language attempts to be colorful, it is oblique. The titles give little hint of the actual content of the various chapters. Finally, there is no easily discernable logical sequence to the chapter headings.

This set of responses was rated as C on all three dimensions.

Chapter 1: The Americans Go to War in Vietnam

Chapter 2: Protesters Start a Major Upheaval in America Over the War in Vietnam

Chapter 3: Americans Soldiers Return to Hatred and Opposition

Chapter 4: The American People are Made Confused about the War by a Media Whirlwind

Chapter 5: Vietnam — the War We Must Try to Understand

The titles in this response are wordy and, in the case of Chapter 4, awkwardly constructed. On the other hand, this students covered all the topics in a logical order. **Consequently, the response was rated as an A in terms of range of topics and logical sequence and C in terms of language.**

Chapter 1: The Real Story Behind Vietnam

Chapter 2: The Proponents and Opponents

Chapter 3: How the Veterans Were Treated

Chapter 4: What Is the Conclusion?

Chapter 5: Is Anyone Going To Figure It All Out?

This response begins well but falters as it proceeds. There is some question whether the title for Chapter 1 reflects the author's tone in the preface. Despite the fact that Chapters 2 and 3 clearly cover most of the topics suggested by the author, the set of titles ignores the role of the media and gives only slight indication that the background of the war will be considered. Although the language is succinct and, in some cases colorful (e.g., proponents and opponents), it lacks clarity in other cases. For example, no indication is given of the differences between Chapters 4 and 5.

Comments

The topic interested the students; their responses indicated that they understood the author's language and had a grasp of the requirements of the task. Even those who did not perform well answered the question with some understanding. In general, students showed more involvement and performed better on this task than on a similar one that asked for an outline of instructions for gravestone rubbing.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Science

Grade 12

Large numbers of dead fish, all white perch, have been found along the shore of the Quabbin Reservoir. Other fish species have not been dying. Give two possible hypotheses explaining these findings. Describe how each of these hypotheses might be tested.

Major Concepts/Abilities Tested

Nature of Science

Scientific Methods

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers

Hypotheses

1st 2nd

** 8% 6% Good hypothesis and appropriate test with controls.

Hypothesis

"The white perch may be reacting to something that they are eating."

How to Test

"First it should be a controlled experiment. A group of white perch must be placed in an aquarium with just fresh water and food. In another tank, water and food from the reservoir should be put in along with the fish. These fish should both be taken from the reservoir. After a few weeks time, the fish should be checked on and the data should be recorded. If the fish in the second tank die, then another test with the same water and different fish should be done."

(The experiment, although convoluted, adequately tests the hypothesis. It could be improved, but the student does address the issue of a controlled experiment and eventually does cover all bases.)

- * 17% 12% **Good hypothesis and appropriate test with some controls.**

Hypothesis

“A corporation might be dumping something into the reservoir that only had an effect on the white perch.”

How to Test

“Test the water for chemicals and if one is found, get a white perch from somewhere else along with other fish from the reservoir and test the perch along with the fish to see if that chemical kills only the perch.”

(There is some control of variables here with the introduction of a white perch from somewhere else, but not enough. The water should also be controlled for by comparing it to other water.)

- ** *good answer*
* *minimally correct answer*

How to Test

“It could be tested by having some test the body of water that this fish is living in to see if there are any chemicals in the water that may affect the fish.”

(This response is very similar to the previous answer but without even the introduction of a control fish.)

- 18% 16% **Other incorrect answer.**

- 9% 18% **Blank**

Comments

This question can be applied to all scientific disciplines, since it examines the concept of scientific inquiry. All scientific inquiry comprises two general aspects: the consideration of evidence based on observations and measurements; and the generation of a hypothesis based on an examination of data, combined with logic and imagination. The hypothesis suggests which data to pay attention to and which data refute it. This question tries to address the issue of selection of data. Here we are concerned with students paying attention to data that will truly measure, test, or give evidence for the hypothesis, we were less concerned about the logic of the hypothesis.

More than half the students were able to form good hypotheses which were logical, reasonable, and could be tested, but faltered when it came to testing the hypothesis. Often they did not control for all the variables or they selected the wrong variables to control for. As evidenced by the results, most often they selected a logical test with no controls. Evidence from this type of test could not conclusively confirm or invalidate a hypothesis.

Incorrect Answers

- 49% 44% **Good hypothesis and logical test but lacking controls.**

Hypothesis

“The water may be polluted with a chemical that only effects that type of fish.”

ON THEIR OWN

Students' Responses to Open-Ended Questions in Science

Grade 12

A local merchant in a town of 30,000 people thinks there may be a relationship between political affiliation (whether people are Democrats or Republicans) and whether they shop at her store. (Out-of-towners rarely shop there.) Describe in detail a study the merchant could conduct to test her hypothesis. Your answer should include a description of the people who would be involved in the study and the information to be collected.

Major Concepts/Abilities Tested

Scientific inquiry

Statistics

Sampling

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers

**** 8%** Identifies the *correct four groups* (town Democrats and Republicans, shoppers Democrats and Republicans), a *sampling plan* (ask the shoppers to indicate their political affiliations) and an *analysis* (compare the results of the shoppers' responses to the affiliation of the town as a whole).

"She could take a random sample of those who shop at her store and see what political party they are affiliated with. She should then compare her results with the percentage of people affiliated with each party in the entire town (this is at the town hall). Comparing these results will enable her to see if her hypothesis is true."

*** 7%** Identifies the *correct questions* AND either a *sampling plan* OR an *analysis*.

"At the entrance to her store, she could have someone simply standing there and

Massachusetts Department of Education

November 1991

Massachusetts
Educational
Assessment
Program

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger and Brenda Thomas.

asking each customer their political affiliations, and then explain what the survey is for. The shop owner should then place another person at a shop on the next block which is similar to the other shop. The 'Questioner' should ask the same question and then compare the results."

(Although this response recognizes the need for a comparison to be made, it does not indicate that the political composition of the town as a whole is necessary for valid comparison.)

*** 6% Identifies the correct groups.**

"She could find out what political affiliation her customers have and the affiliation of all the other citizens. A simple questionnaire could be handed out to her customers to see if their shopping there had any influence on their political choice."

(This response does not address the relationship between the citizens of the town and the shoppers.)

** good answer

* minimally correct answer

Grateful Dead T-shirts and buying hot dogs, they're probably Democrats."

(This method only measures the prejudices of the observer.)

"The merchant could offer bumper stickers for each party. By doing this, she could see how many bumper stickers for Democrats and Republicans are taken."

(This method is more reliable in discovering political affiliation than the previous one since it requires that customers self-identify with a particular party. Still, it does not address the relationship between the shoppers and the town.)

13% Other incorrect response.

13% Blank

Incorrect Answers

40% Identifies the correct information but the wrong group (e.g. just ask shoppers or just the general population). It does not address all four groups.

"She could set up a survey which is to be filled out by every customer that goes to her store. It should be comprised of questions in relationship to which political party they belong to, whether or not they are local residents. This survey should give her all the information she needs."

13% Identifies the shoppers' affiliations indirectly.

"If they're wearing three-piece suits and buying caviar, they're probably Republicans. If they're wearing

Comments

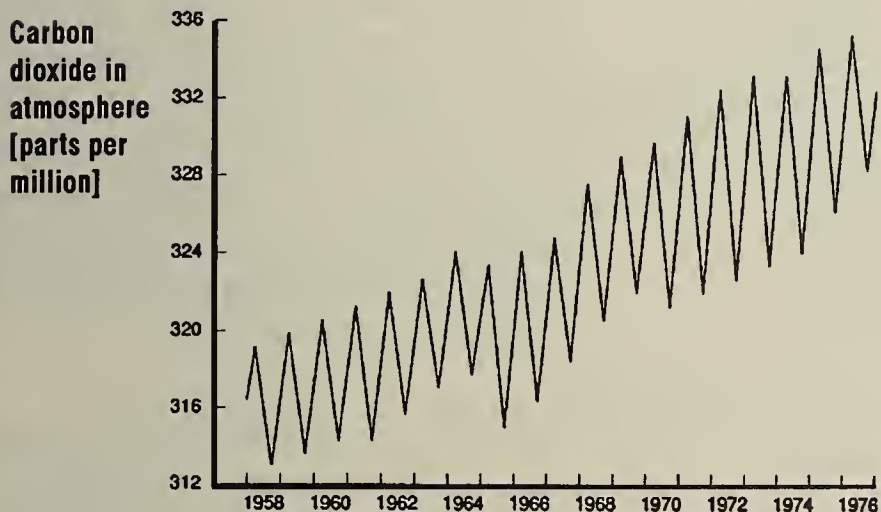
Although many students were able to identify the correct question e.g., "Are you a Democrat or a Republican?", most were not able to identify the correct groups, e.g., the town as a whole and the shoppers in the store. The most common response was to ask the shoppers whether they were Democrats or Republicans. This implies an inability to see the larger picture. They can make the very obvious connection between the shoppers and the store, but cannot see that, if the percentages of Democrats and Republicans shopping at the store are to have any significance, they must be compared with the percentages of Democrats and Republicans in the town.

The parenthetical phrase in the question "(Out-of-towners rarely shop there.)" confused students immensely. Many suggested strategies for generating business with out-of-towners; others simply suggested ways of generating business and did not answer the question at all.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Science

Grade 12



The graph above shows the concentration of carbon dioxide in the atmosphere over a period of almost two decades. What reason or reasons can you give for the general increase in carbon dioxide over the time period from 1958 to 1976?

Major Concepts/Abilities Tested

Environmental Issues

Impact of Technology

Reading and Interpreting Graphs

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers

*** 24% Two or more of the following:
increase in automobiles, increase in industry, increase in population, increase in use of fossil fuels, decrease in forests or any other reasonable answer with elaboration that shows some understanding of the causes.

or

One reason with knowledgeable elaboration. Student is able to iden-

tify a cause and discuss how it is related to the effect (discusses interaction of the two).

“The population of the earth has increased considerably between 1958-1976 and human beings are great producers of carbon dioxide. The vegetation and forestation of the earth also considerably decreased 1958-1976 and plants are great consumers of carbon dioxide. The number of cars and use of the car also increased during these years, and car exhaust added to the carbon dioxide in the atmosphere. The continued burning of fossil fuels in other areas also contributed to the increase in carbon dioxide 1958-1976.”

(This response gives many reasons for the increase in carbon dioxide.)

“The general increase in the concentration of carbon dioxide gas in the atmosphere over the past two decades has been a result of the development of new technology. That is, there are more cars on the road than before spewing their poisonous gases into our air. Similarly, new technology has created a greater need for industrial power — often obtained by burning coal which releases carbon dioxide. New technology has produced many new manufactured items which sometimes require, as do plastics, the combustion of certain substances in their creation.”

(This response deals with one large category, “new technology,” in more depth than the first answer.)

**** 32% One or more of the above reasons with some elaboration.**

“Because more people are buying and driving cars and, of course, cars blow out a lot of carbon dioxide when they burn fuel.”

(Although this response gives a cause and a reason for that cause, the answer is slight.)

*** 19% Minimally correct; identifies a cause.**

“The reason for the increase in carbon dioxide from 1958 to 1976 is from the use of fossil fuels.”

(This answer does not detail why the use of fossil fuels causes an increase in carbon dioxide.)

***** good answer**

**** correct answer**

*** minimally correct answer**

Incorrect Answer

14% Incorrect explanation.

“The increase of parts per million of carbon dioxide could have been caused by the disintegrating of the ozone layer. The ozone layer exists somewhere between the surface of the earth and the sun. It essentially blocks the harmful ultraviolet rays of the sun. It is gradually disintegrating, scientists say, because of excess use of aerosol sprays such as hairspray and deodorant sprays. Thus the ozone layer can no longer absorb (as much as it used to) the carbon dioxide in the atmosphere. This effect is known as global warming.”

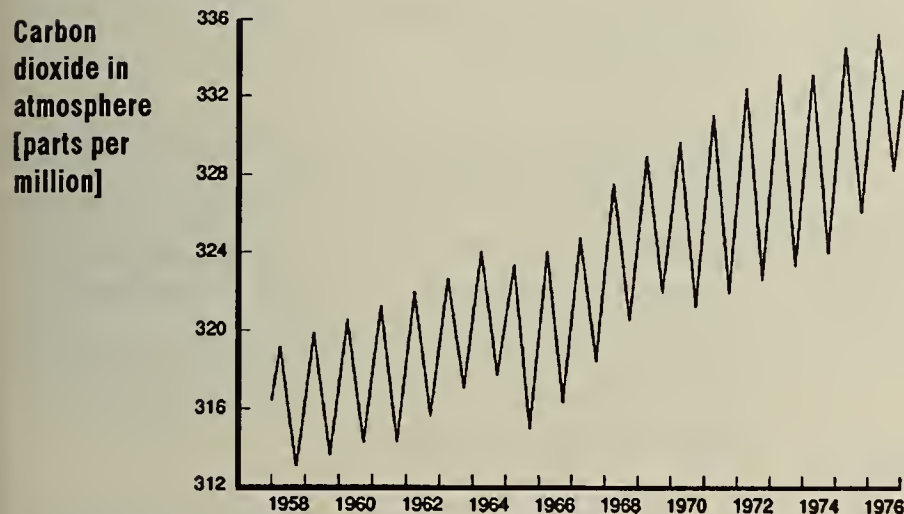
(Although the definition of the ozone layer is partially correct, the disintegration of the ozone layer did not cause the increase in carbon dioxide.)

“Plants give off carbon dioxide into the atmosphere so there has been more plants and trees in the land giving us more carbon dioxide.”

10% Blank

Comments

The responses to this question reflect the fact that schools are placing a greater emphasis on environmental issues, with three-fourths of the students able to give at least a minimally correct response to this question. From this starting point, students can begin to see the relevance of science in their day to day lives. Many students and teachers consider science as strictly for scientists, when in fact, the reality of science is intimately connected with all aspects of life.



What reason or reasons can you give for the fluctuations (ups and downs) that appear in the graph?

Major Concepts/Abilities Tested

Reading and interpreting graphs

Environmental issues

Correct Answers

**** 17% Seasonal fluctuations mentioning plants and/or increase use of fossil fuels in the winter.**

“The fluctuations in the graph vary between summer and winter. In the winter, many places are heated, increasing carbon dioxide levels. In the summer, the need for heating decreases and so does the level of carbon dioxide. In the summer, the leaves on plants change carbon dioxide to oxygen. The peak of the carbon dioxide levels occurs in the winter.”

*** 7% Seasonal with no explanation or partial explanation.**

“The reason on the whole why the graph is going up is obvious, because of the increase in carbon dioxide. The reason for the fluctuation is because, not only is the carbon dioxide changing during the years but during different seasons of the year, so it is constantly changing all year round.”

(Unlike the first answer, this student does not explain why the seasons cause the fluctuations.)

*** 4% Measurement error or needs more information.**

“The main reason I feel there are fluctuations is because as with data from all experiments, this data is subject to experimental error. So at different times, factors could have come into play causing incorrect results. Thus the fluctuations in the graph.”

(Although this answer displays an astute sense of science, the student misinterprets the effect of measurement error because with such a large quan-

tity, the effect would be slight and the effect would not have the regularity that is displayed in the graph.)

** *good answer*

* *acceptable answer*

Incorrect Answers

20% Does not recognize the regular fluctuation; instead discusses the overall trendline.

“I do not know exactly why, but if you look at the dates you can see there was a war being fought in some of the higher numbers of the graph like after 1966 when the Vietnam War was going on. We have also revolutionized new fuel burning products and we have done nuclear testing. These might be the reasons for the fluctuations in the graph.”

(It's clear that the student does not address the regular fluctuations, but rather looks at the trendline and tries to attach historical events to the graph.)

“The main fluctuations of the graph might have occurred because of the wars where many people were killed. That would put less carbon dioxide into the air so the amount would go down. The reason it goes up after the down spell is because of the baby boom period after the war. As of right now, that happens to occupy the biggest age range percent of the people in the population today.”

(Apparently this student took some information from a social studies graph on population by age groups, which appeared on another page of this test. Although the idea of synthesizing various pieces of information is generally useful and appropriate, in this particular situation the student selected the wrong data to use.)

25% Other incorrect answer.

“The reasons for the fluctuation is that when there is a great deal of rain and snow one year this will reduce the amount of carbon dioxide in the air.”

“The reasons for the fluctuations were probably that the economy was down and the price of gas went up and therefore people's cars were being used less. That is what I would think.”

27% Blank

Comments

This question focuses on two separate aspects: understanding the data presented; and applying the data to the real world.

The first aspect simply deals with the individual details in the graph. Did students read the legends and any comments about the graph? Were they able to understand the relationship in the graphs as relationships? In other words, did they understand the data simply as data? We wanted students to notice the fluctuations and consider their significance.

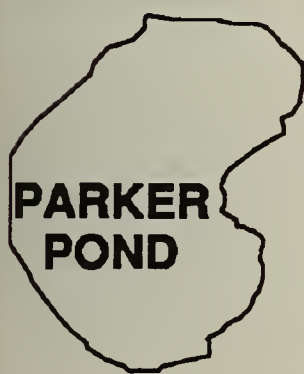
The second aspect focuses on whether they were able to correctly relate the data to the actual events represented. In other words, the students who answered that the fluctuations were seasonal paid attention to the years listed on the horizontal axis and the space between the fluctuations. By looking at both these facts, they noted the fact that the fluctuations occurred regularly within one-year periods.

This proved to be an extremely difficult task for the majority of students. Over one quarter of the students did not even attempt the question, and more than half of those responding failed to recognize that the regularity of the fluctuations suggested a naturally occurring event.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Math

Grade 12



Theo wants to find out which pond covers the larger area. He does not need to know the two areas, just which is larger. Theo claims that all he has to do is measure the distance around each pond to find out what he wants. Will his method work? Make a convincing argument for your answer.

Major Concepts/Abilities Tested

Understanding of area/perimeter

Ability to communicate mathematically

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers

*** 9% No, discusses difference between perimeter and area and gives counter-example.

"This method will not work. What Theo is measuring is circumference, which is different from surface area. One pond may have a circumference with many bends and turns increasing the circumference, whereas the other may be more spherical with few turns, lowering the circumference. But the surface area of the pond with less circumference may be much larger than the surface area of the pond with many

bends. The two dimensions, area and circumference, are different and this should not be compared.”

(This student answered the question by producing an argument why the same perimeter does not necessarily imply the same area. The student appears to recognize the requirements of a convincing argument by giving a counter-example.)

**** 8% No, states difference but no argument.**

“Theo’s argument would not work. If he measured the outside of the pond by walking around it, he would get the perimeter of the ponds, not the area that they cover — which is what he needs.”

(Merely defining the terms “perimeter” and “area” does not constitute an effective argument against a proposition.)

*** 17% No, minimal response without explanation.**

“Wrong. Because he wants to find which covers a larger area.”

*** *good answer*

** *correct answer*

* *minimally correct answer*

Incorrect Answers

13% No, explanation unclear or incorrect.

“No, he not only needs to measure the distance around each pond but the diameter, and the water depth. After all three measurements, he would be able to make a more accurate answer.”

(This student interprets Theo’s argument in terms of volume, despite the fact that “covering” is a terms that refers to area, not volume.)

“He could simply measure both ponds’ width and length and determine that way which is larger.”

(This student is typical of many who responded with an automatic application of a formula to the word “area.”)

37% Yes

“Yes, because if you measure the distance around both of the ponds and compare them, then the larger measurement will cover the larger area.”

“The distance around each pond measures how large of an area it covers. So that would be the logical thing to do.”

10% Blank

Comments

Vocabulary is usually cited as the reason for the notoriously poor results on area problems. Undoubtedly, a confusion in the meaning of “area” and “perimeter” accounts for some errors, but the results above point to a more profound source. When given a problem that tapped their conceptual understanding, only a third of the students were able to recognize the fact that “area” implies something more than a linear measurement. Furthermore, less than 10% were able to produce a convincing argument for their belief. The responses to this question clearly illustrate the need for discussion in mathematics. Unless students are able to examine and explore different instances of a concept, they will be reduced to a mindless application of barely remembered formulas.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Math

Grade 12

Roger says that raising the score on a high-scoring test paper would raise the class average on the test more than raising the score on a low-scoring paper by the same amount. Is Roger right or wrong? Use the space below to explain or prove your answer to someone who does not agree with you.

Major Concepts/Abilities Tested

Statistics

Reasoning

Communication

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers

**** 20% Roger is wrong with a good explanation.**

"He's wrong. As long as the number of points added was the same, it wouldn't matter. To get the class average they have to add all the scores up and then divide by the number of people in the class. The total number of all the scores will be the same no matter where the points were added."

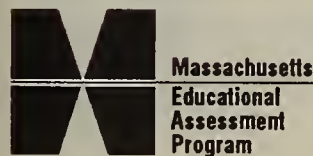
"Raising one score or another by the same amount would increase the average equally. Since average is a computation of the total points in the test, the same average will be attained."

**** 16% Roger is wrong with a good example.**

"Either way you add the points you are going to end up with the same amount. For example, Roger's grades on two quizzes are 80 and a 20. If the teacher adds 20 points to the low grade, the grades would be an 80 and 40 for an average of 60

Massachusetts Department of Education

November 1991



In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger and Brenda Thomas.

percent. If you add 20 points to the higher grade, you get 100 and 20 for grades. Roger's average would still be 60 percent. I feel Roger is wrong. Whichever way you add the points the final outcome will be the same."

*** 15% Roger is wrong without justification or unclear justification.**

"Wrong because the test average would be the same because they are raising the score by the same amount."

** *good answer*

* *minimally correct answer*

Incorrect Answers

There was a wide range among the incorrect responses. Generally they broke down into the following categories:

5% High scoring papers should count more.

"If you have a 50 on a test and the teacher gives everybody ten more points you would still have a failing mark. But if you have an 80 on a test and the teacher gives everyone ten more points you would have a 90 from a B to an A, but in the other case you would still have an F."

"Roger is right, because if you take a paper with a high score and you raise the score on the high scoring papers you will bring the average grade up, but if you take a low score and make it higher it really won't affect the class average much."

7% Low scoring papers should count more.

"Roger is wrong because if you raise the score on a high-scoring test paper, it will make the rest of the tests lower than if you raise the score on a lower scoring test. If the lowest mark in the class is raised, then all the other tests have to be raised also, which can only make the class average higher."

5% Cites motivational reasons.

"I think he should raise the score more on the low-scoring papers to give the students who need it more of a chance."

"I don't agree with Roger. You shouldn't raise anybody's paper. Because you shouldn't praise those who learn quicker. You should be fair about education and grades. Don't forget we are the children of the future and we all should get the grades we deserve."

19% Other

"Roger is right, it is harder to raise the higher scoring test paper than the low scoring paper."

"Depends on what kind of averages you are talking about, mode, mean or medium."

11% Blank

Comments

A variation of this question was asked in 1988. On that occasion, a quarter of the students did not attempt the question, while 48 percent were able to supply a correct response, with four percent constructing an algebraic proof. The present version elicited far fewer satisfactory responses, but more students attempted it. Its context (the raising and lowering of student scores) may have seemed more accessible and certainly diverted some students into believing that the issue was one of fairness rather than of statistics. Nevertheless, we believe that the comments written in 1988 are equally applicable to these results:

"It was obvious from their discussion that many students had only a shaky understanding of what an average means. Many more had little understanding of the hallmarks of a mathematical argument."

ON THEIR OWN

Students' Responses to Open-Ended Questions in Math

Grade 12

In a recent survey, Americans were asked about ownership of firearms. The findings of the study were that

* 25 percent of American families have at least one handgun;

* 25 percent of American families have at least one rifle; and

* 10 percent of American families have at least one automatic rifle.

A reporter used the following headline on an article she wrote about the study:

MAJORITY OF AMERICAN FAMILIES OWN FIREARMS

Should the reporter's editor accept the headline as it is? Why or why not?

Major Concepts/Abilities Tested

Statistics

Percentages

Correct Answers

* 25% No, with correct reason.

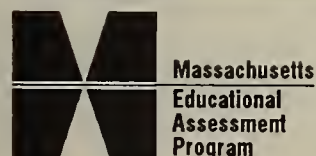
"The headline is probably false since a family who owns a handgun might also own a rifle or automatic rifle. The headline might read: Many American families own firearms."

"He should not. He has no proof that it is a majority of Americans. It is more likely that the Americans owning a firearm have more than one in their possession. Therefore the title must be renamed."

(These students correctly understood that the sets of persons in each category of gun ownership could intersect. There is no way of knowing this from the survey.)

Massachusetts Department of Education

November 1991



In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger and Brenda Thomas.

Incorrect Answers

34% No, with incorrect reasoning.

"No they shouldn't print it because only 25% of families own a gun. 25% isn't a majority."

(This student does not appear to recognize that rifles can be classified as firearms.)

"No this should not be used as a headline, a survey is only a sampling of the population, therefore not everyone answered. So we do not know if the majority of American families own guns."

"No because it doesn't say how many people were asked out of how many people."

(The first student was not willing to accept the results of a survey. If reference had been made to the possibility of a sampling error, the response might have been more convincing. As it stands, it seems to imply that credible results can be obtained only by asking each individual. The second student did not seem to understand the meaning of percentages as summary statistics for "how many people were asked out of how many people.")

"I don't think that the reporter's editor should accept the headline. I feel that it is not catchy enough. The headline is exaggerated. The majority of Americans families don't own firearms."

(This response is typical of those students who objected to the word "majority.")

35% Yes

"The reporter's editor should accept the headline, because more than half of American families own firearms. I would consider 60% of American families owning guns a lot of people and I find nothing wrong with the title."

(Almost all of the students who answered that the headline should be accepted did not recognize the possibility that the same families could be represen-

ted in more than one category.)

8% Other incorrect answer.

"No it should read 'Why are the Families of America Feeling so Insecure?' and then state the statistics."

"No, it should be something to catch the readers or viewers eye. Something like, A Nation at Arms."

(These students appeared to see the question as a stylistic task.)

"Yes, he should accept this headline, because there is freedom of speech and the freedom of press."

"No, because people of the U.S. will think that it is right to own and use them at their own will."

(Students who responded in this manner, possibly misunderstanding the purpose of the task, did not address the accuracy of the headline itself.)

3% Blank/I don't know

Comments

An important practical skill for any citizen is the ability to understand the statistics behind the headlines. The responses to this question suggest that only a quarter of twelfth graders are capable of doing that. Many merely added up the percentages given, without recognizing the possibility that the same families could belong to more than one category. Beyond this, however, the question elicited a wide range of responses which gave an insight into students' understanding of the role of sampling, the function of percentages, and the meaning of the word "majority." It appears that these notions, which are used to support a great deal of political and economic discussion, are not well understood by many twelfth graders.

NRA is very strong and Congress is afraid of them.”

**** 13% The answer contains two of the above elements.**

“Congress wants to, or is debating whether to control the use and purchase of handguns by putting a seven day delay on their purchase. However, the National Rifle Association is very much against the control of handguns, as they consider the right to bear arms a constitutional guarantee.”

(This answer is very similar to the first example; however it does not address the weakness of Congress in the face of pressure from the NRA.)

*** 17% Answer contains one of the above elements.**

“The National Rifle Association doesn’t want a seven-day wait period for purchasing handguns.”

(This answer is marginally correct. It doesn’t contain any wrong information; there is not enough of it. By identifying the NRA, the student indicates he has some sense of the cartoon’s intent.)

******* *Very good answer*

****** *Good answer*

***** *Minimally acceptable answer*

Incorrect Answers

26% Correct information but doesn’t address the question.

“The National Rifle Association fights for what’s in the constitution, the right to bear arms. And they will fight for that right. We shouldn’t have to change what’s in the constitution.”

(Although what the student says about the National Rifle Association may be true, this response does not discuss the main idea of the cartoon.)

17% Response indicates lack of prior knowledge or literal interpretation of characters.

“The man being killed by the NRA wants a seven day waiting period so he can escape.”

12% Other incorrect answer

“No guns should be sold to anyone who has not a permit. Abolish guns is what they are trying to say.”

9% Blank

Comments

Although students are exposed to cartoons from a very early age, many students are unable to grasp the idea of humor in the form of satire as a means of revealing some societal defect. Very few students were able to identify more than two elements of the cartoon. Because Congress and the NRA are labelled, most twelfth graders had no problems identifying the main characters of the cartoon, despite some imaginative answers (Nancy Reagan Association and National Reclamation Agency).

When they had to interpret the meaning, they often fell short, being unable to move from the representational to the real. Students appear to lack the ability to take the very simple situation illustrated by the cartoon and attach a larger significance to it. Just over one third of the students responded with some degree of correctness, indicating that they are not able to make the leap from the literal to the figurative. Because of their tendency toward literal interpretation, the use of irony and exaggeration only exacerbates students’ difficulties.

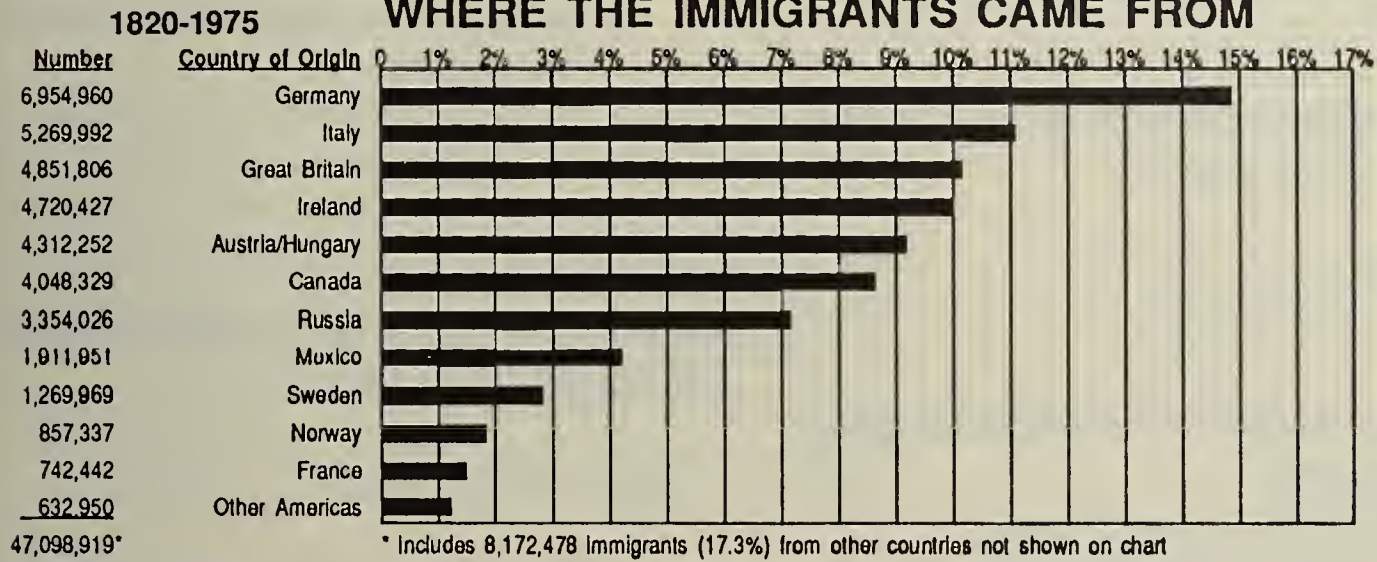
This question also evaluated prior knowledge, which often was incorrect. Many seniors believed that the National Rifle Association was opposed to guns; others felt they sold guns.

ON THEIR OWN

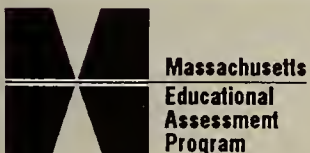
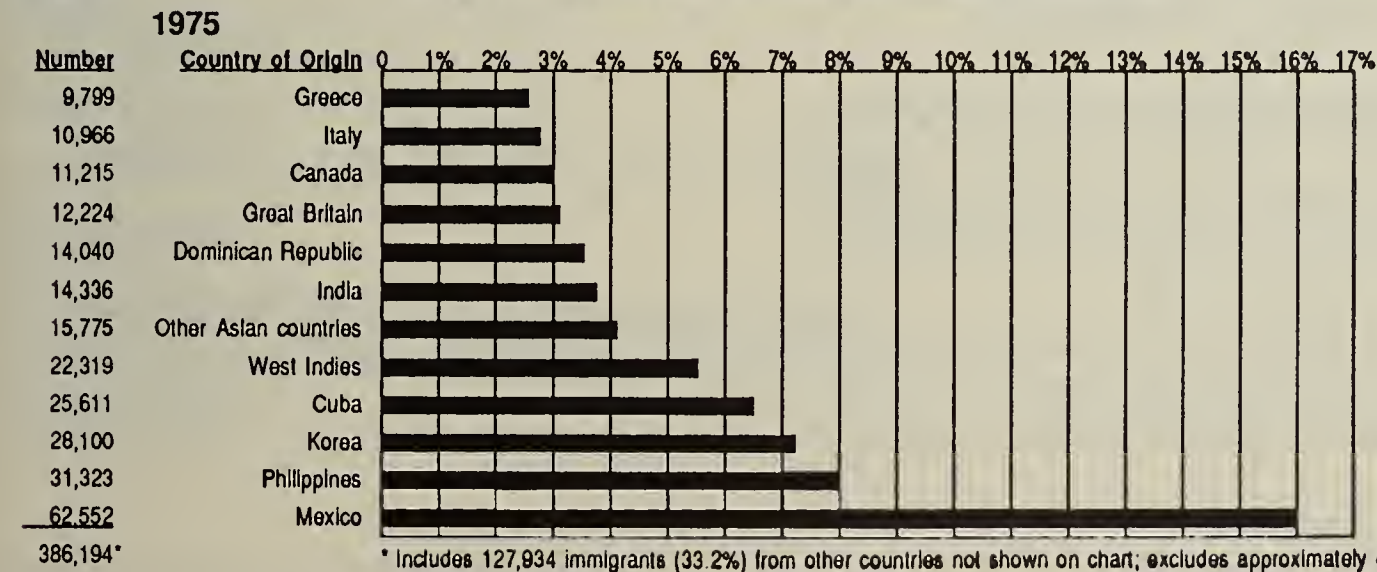
Students' Responses to Open-Ended Questions in Social Studies

Grade 12

WHERE THE IMMIGRANTS CAME FROM



From *Sociology: The Search for Social Patterns* by Larry F. Krieger and Ira Peck. Copyright © 1980 by Scholarship Inc. Reprinted with permission.



Massachusetts Department of Education

November 1991

In the spring of 1990 over 9,000 fourth, eighth, and twelfth grade students were assessed using open-ended mathematical, scientific, social studies, and reading concepts. This series of reports describes and discusses the results of these assessments. Prepared by Brenda Thomas and Elizabeth Badger.

The numbers of immigrants are given to the left of each graph. Why are there more immigrants listed on the top graph than on the bottom one?

Major Concepts/Abilities Tested

Interpretation of Graphs

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers

- * 35% Recognizes that the number of years is different.

“The top graph spans 155 years, the bottom graph is one year.”

- * *correct answer*

Incorrect Answers

- 60% Incorrect answer

“More immigrants are listed on the top graph because the top graph is before the U.S. put low caps on immigration. Before the U.S. wanted immigrants to fill the U.S., then they became full and had to set limits.”

“The illegal immigrants were excluded from the bottom chart. Also 33.2% from other countries are not shown and in the top one only 17.3% are not shown.”

- 5% Blank

In the top graph the number of immigrants from Mexico is almost two million (1,911,951). In the bottom graph, only 62,552 are listed. Why is it that the line for Mexico is so much longer in the bottom graph than in the top one?

Correct Answers

- * 44% Answer shows the students understands that the number is a percentage.

“The lines show percentages of total immigration. In the bottom graph, a majority (in this case 16%) of the immigrants are from Mexico.”

- * *correct answer*

Incorrect Answers

- 46% Incorrect answer

“Because in the bottom graph it says that there are approximately 400,000 aliens. I would imagine that at least one half or maybe all of them are from Mexico.”

“Because it wasn’t until 1975 that we let them into our country, before that we limited the amount.”

- 10% Blank

Comments

The purpose of these two questions was to lead students to the appropriate reading of the graph. We felt that the graphs might be too difficult for students without some guidance. In the first question, students were required to recognize that the graphs dealt with different time spans; in the second, it was necessary for students to understand that the lines were percentages and not numbers.

In order to answer these two questions correctly, students had to examine the graph carefully and draw very obvious conclusions. The answers were clearly embedded in the graph. If the first skill to be used in reading a graph is correct observation of the data, then less than half the students were able to read the data correctly. Interestingly, many students mentioned the fact that 400,000 illegal immigrants were excluded, indicating that those students had read all the fine print but had not selected the appropriate facts for answering the question.

In both questions, students did not perform up to our expectations, with less than half of the students answering either question correctly. It may be that, considering the complexity of the other open-ended questions, students felt that the correct answers were too simple and, therefore, attempted to elaborate where there was no need. In response to the first question, many students attempted to explain why immigration had decreased, based on their misinterpretation of the graph. In the second question, students attempted to explain the increase of immigrants from Mexico.

The previous two questions were asked to prepare students for the next question.

What changes in the U.S. immigration pattern do the graphs show?

Major Concepts/Abilities Tested

Interpretation of graphs

Generalizing from data

Correct Answers

**** 28% Shift in immigration with generalization about trends.**

“The chart shows that between 1820 and 1975, immigration from European countries was high, but more recently, the majority of the immigrants come from Asian and Spanish countries.”

*** 11% Shift in immigration with no mention of countries.**

“The countries where the least amount of people came in 1820-1975 has reversed in 1975. In 1975, the least became the most.”

*** 8% Shift in immigration mentioning specific countries, not areas.**

“Fewer people are coming from Germany and Italy and more are coming from Mexico.”

Incorrect Answers

4% Increase in number.

“That an abundance of people are starting to come to the United States more and more each year.”

28% Decrease in number.

“Not as many people are immigrating now as there were before.”

13% Other incorrect answer

“It shows that a lot more people are coming to this country illegally because in the

bottom graph they excluded 400,000 illegal immigrants and not in the top.”

“That more and more people are coming to the United States illegally and not through the immigration process. Maybe the U.S. has some tight quotas.”

8% Blank

** *good answer*

* *minimally correct answer*

Comments

We asked this question to examine students' ability to generalize about immigration trends based on the information given in the graphs. We were not looking for specifics, but the recognition of trends. Students were expected to evaluate information about specific countries and to draw conclusions based on the information presented.

Just about half the students were able to recognize and describe a pattern in the immigration rates based on the evidence given in the graphs. Although calculation reveals that the number of immigrants has increased, very few students recognized this. Of those who mentioned the decrease, many tried to explain the changes as the result of tighter immigration quotas. In other words, instead of identifying the changes that occurred they attempted to answer *why* certain changes (either real or perceived) had occurred.

We can see from the responses that students do not read graphs carefully. Although students use graphs and charts at early ages, they don't seem to apply their graph reading skills in a consistent manner.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Social Studies

Grade 12

Historically, the United States has had a relatively open policy toward immigration. Discuss some important ways this policy has influenced American society . . . economically, politically, socially.

Major Concepts/Abilities Tested

Approaches to history
Multiple points of view
Knowledge of immigration
Relationship of past to present

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers

** Well-developed answer; each statement gives appropriate evidence and represents a balanced response.

** 13% **Economically**

"The United States' open immigration policy led to a bigger workforce at the turn of the century. Many immigrants came to America to escape oppression and to seek employment. The industrial revolution had come and the factory system thrived on having this new source of cheap labor."

** 11% **Politically**

"The United States' open immigration policy has led to its having better relations with other countries than it might have had. Occasionally, however, the U.S. policy toward immigration has caused some tension — Chinese Exclusion Acts of the late 1800s/early 1900s for example or the

reluctance of the United States to accept the flood of German Jews in the 1930s and 1940s.”

**** 13% Socially**

“Having an open immigration policy has led to the U.S. having a relatively more tolerant and integrated society. A great mix of nationalities populate the United States and make it a very culturally rich country. Naturally there have been problems such as the discrimination against the Irish in Boston, but on the whole, immigrants have helped the country with different foods, clothing, music and styles.”

*** Acceptable answer; some development to the response.**

*** 33% Economically**

“Because immigration took place, many jobs were filled and little money was needed to pay workers since immigrants demanded such little pay. Companies were allowed to save and to continue to develop into industries.”

(This is similar to the economic response in the first category, but does not include the historical dimension.)

*** 27% Politically**

“The influx of immigrants had allowed many different viewpoints by allowing a greater number of options and the opportunity for the American people to decide which one is right.”

*** 31% Socially**

“We are expanding our culture by involving new cultures in our society. We are learning from them just as they are learning from us.”

(Although they are thoughtful, the two responses above are too generalized to be considered well-developed.)

Incorrect Answers

Statement with no development.

13% Economically

“A lot of low paying jobs have been taken by immigrants who will work for less money when they get here.”

17% Politically

“I believe that politically it has done some good because we believe that Americans are not made of only one race and are made of many different races and cultures, thus making us the great nation that we are.”

17% Socially

“Immigrants may have caused some problems socially that have been brought on through no fault of their own. I speak about discrimination that ‘old comers’ feel for the ‘newcomers.’ The difference, of course, being that the old comers were here before the newcomers.”

(Although there appears to be an attempt at elaboration in these responses, the elaboration only tends to restate the thesis statement.)

Answer only discusses negative aspects of influence of immigration.

22% Economically

“Immigrants work in low-income jobs. However, these jobs are taken from Native Americans so the rate of unemployment will increase.”

(It is noteworthy that almost one-fourth of the seniors responded only negatively

when asked to discuss the economic influence of immigrants on the United States. The depressed economy seems to generate scapegoats. This student's use of the term Native American to mean native-born people of immigrant descent is typical of many students.)

8% Politically

"I feel that the political aspect of immigration is wrong. It's great that people can come to the United States and share our liberated freedom, but not so in the political field. Politics in our government should be run and voted upon by Americans only. Immigrants don't know enough about America to run it."

(This response reflects many that fell into this category. Students seem to feel that only those born in the United States can know about it.)

15% Socially

"I think it has done some good and some harm because we have to give immigrants a chance in improving and changing their lifestyle. It has also done some harm because now we are having racial problems. Americans are blaming the immigrants for all their problems of poverty, drugs, and even death. We have problems with Blacks against whites, Hispanics against Chinese, and everyone trying to create their own separate community because they have separate backgrounds."

(Although this student comments that immigration "has done some good" the only elaboration in the response deals with negative aspects of immigration.)

Does not address the specific area called for.

5% Economically

7% Politically

6% Socially

"Immigrants have been able to adapt easily to American culture. Jobs, housing, and communities have been kept open because of interaction between various types of people. They come to this country and we teach them our way of life."

(This student comments on how the United States has affected the immigrant rather than the effect of immigration on the United States.)

Open door immigration policy has not influenced that area of American society.

0% Economically

6% Politically

"As of today, there have been no immigrants who have run for office."

(It is only in the political category that students responded in this manner. Rather than examine one of the historical aspects of the political influence of immigrants, such as the rise of machine politics, many students simply considered that only a person born in America can run for president. There appears to be no historical consideration to this type of response.)

1% Socially

Other incorrect

7% Economically

"We have been able to trade freely with other nations. We are able to compensate for areas in which we are lacking (the growth of certain crops) by taking advantage of other countries abundance."

8% Politically

"We have not experienced a great deal of conflict with travelling through nations or passing laws between nations. Negotiations between countries have been ameliorated by open policy."

4% Socially

"In a way, the United States wants everybody to get along with everybody and sometimes it is the people in power who make everyone hate each other."

Blank

8% Economically

(Students had the most to say about the economic influence of immigrants, although what they say is often entirely negative.)

18% Politically

13% Socially

** *good answer*

* *acceptable answer*

General Attitude

Does the student seem to display an unbiased attitude toward immigrants?

77% Yes

9% No

13% Can't tell

Comments

Students were asked to generalize about the immigration experience as a whole, within the various contexts of economics, politics, and culture, and relying on their knowledge of historical and current immigrant experience.

Most students discussed the effects of immigration only on contemporary America and did not account for historical influences or aspects. Because so few students dealt with history, there is little sense that seniors view the past as having any significance to the present. For example, a large number of students focussed on African-American immigrants. While it is true that African-Americans are not native to America and that many have immigrated in the traditional sense, many do not have the same experience as European settlers to this country. Because African-Americans were not willing participants in the immigration process, their immigrant experience cannot be considered similar to the family from Germany who came to America in the late 19th century in hopes of a better life. Yet students did not focus on the historical aspect of African-American immigration; rather, they considered it to be the same as European immigration or similar to that of their own families. This inability to consider multiple perspectives often leads to simplistic understanding.

ON THEIR OWN

Students' Responses to Open-Ended Questions in Science

Grades 8 & 12

The moon orbits the earth. Explain why the moon doesn't fall to earth.

Major Concepts/Abilities Tested

Physical science

Motion

Gravity/inertia

The percentage of student responses and appropriate examples are listed under each of the categories below.

Correct Answers

Gr8 Gr12

*** 0% 3% **Mention of vectors and resultant motions caused by inertia and gravity with explanation or drawing.**

"The moon has both angular and linear velocities. While the moon orbits the earth, the earth pulls down on it with the force of gravity. The moon fights the pull of gravity and wants to fly away from the earth. The moon stays in orbit and moves with a velocity tangent to the elliptical orbit (vector diagram included)." (grade 12)

** 6% 12% **Answer discusses two dimensional motion resulting from inertia and gravity.**

"The moon is kept in orbit by the gravity of the earth and the moon wanting to go off in a straight line.

The two opposing forces keep it in a circle around the earth.” (grade 8)

(The student mentions two separate motions but does not indicate vectors. This response is weaker than one predicated on vector analysis but it does indicate a high degree of comprehension.)

* 10% 11% **Correctly describes either the effect of gravitational force or the inertial motion of the moon.**

“The moon is extremely distant from the earth thus the force of gravity between them is weak in comparison to the force between the earth and ourselves. The force of the gravitational pull is strong enough to keep it in orbit but not strong enough to pull it down.” (grade 12)

“The moon maintains its orbit because its mass is so great it isn’t pulled in. Plus the circular orbit keeps it moving in space.” (grade 8)

(Although this example contains some flaws, e.g., a large mass would increase gravitational attraction and orbit does not cause the movement, the student indicates an understanding that the moon’s orbit is dependent on one of the two factors—the moon’s inertia.)

- *** *very good answer*
- ** *good answer*
- * *minimally correct answer*

Incorrect Answers

13% 12% **Identifies TWO OPPOSING FORCES but not correctly.**

“The moon remains away from the earth because the outer force of the universe acts upon it. The earth’s gravity balances this force so that the moon remains in orbit.” (grade 12)

36% 26% **GRAVITY with incorrect explanation of this one force.**

“The moon does not fall to earth because of gravity, an electromagnetic field which balances the universe.” (grade 12)

5% 9% **The moon is in its orbit/place/position.**

12% 12% **Response suggests there is no gravity in space.**

12% 11% **Other incorrect response.**

5% 7% **Blank**

Comments

The lack of growth between grades eight and twelve is particularly noticeable in the responses to this question. Those students who answered gravity were on the right track; they seem to have some inkling of the relationship between gravity and motion, but could not reason the relationship between the two. Instead they arrived at a convenient explanation of why the moon didn’t fall to earth due to gravitational attraction. This corresponds to research indicating that students often lack a real understanding of scientific concepts. The meaning of a term becomes attached to the particular context in which it is taught and does not transfer to other relevant situations. The contradictions that occur are often unrecognized by the student or teacher.

The principle entailed in the interaction of gravity and inertia is universal. It can be emphasized in many forms across the teaching of science, so that students who understand the principle of falling bodies on earth can apply that principle to the earth and the moon. Responses to this question imply that students do not recognize the application of the principle in this context.

Grade Twelve

